Inexact Newton Dogleg Methods

Homer Walker
Mathematical Sciences Department
Worcester Polytechnic Institute
August 8, 2003

Joint work with Roger Pawlowski (SNL), J. N. Shadid (SNL), J. P. Simonis (WPI).

Nonlinear problem: $F(x) = 0, \quad F : \mathbb{R}^n \to \mathbb{R}^n.$

Start with classical . . .

Newton's Method:

Given an initial x.

Iterate:

Solve $F'(x)s = -F(x)$.

Update $x \leftarrow x + s$.

Globalizations.

Idea: Repeat as necessary ...

- *Test* a step for acceptable progress.
- If unacceptable, *modify* it and test again.

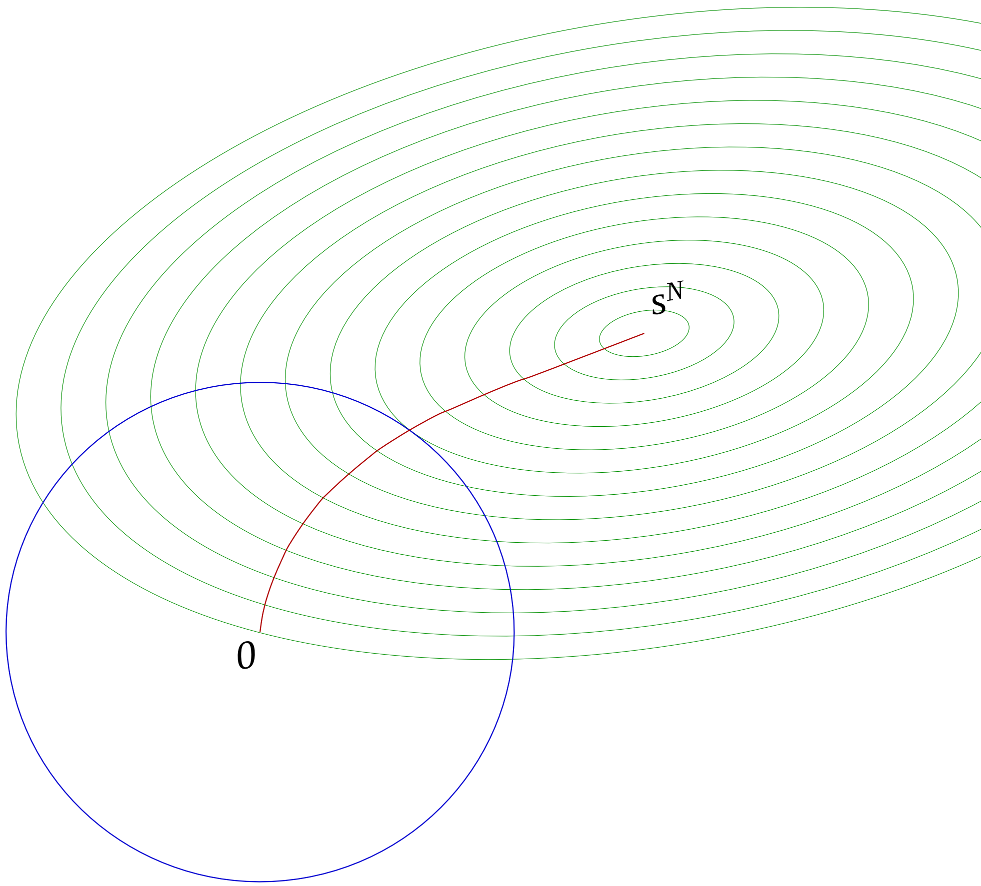
Major approaches:

- *Backtracking* (linesearch, damping).
- *Trust region*.

Trust region globalization.

- $s = \arg\min_{\|w\| \le \delta} \|F(u) + F'(u)\, w\|$.

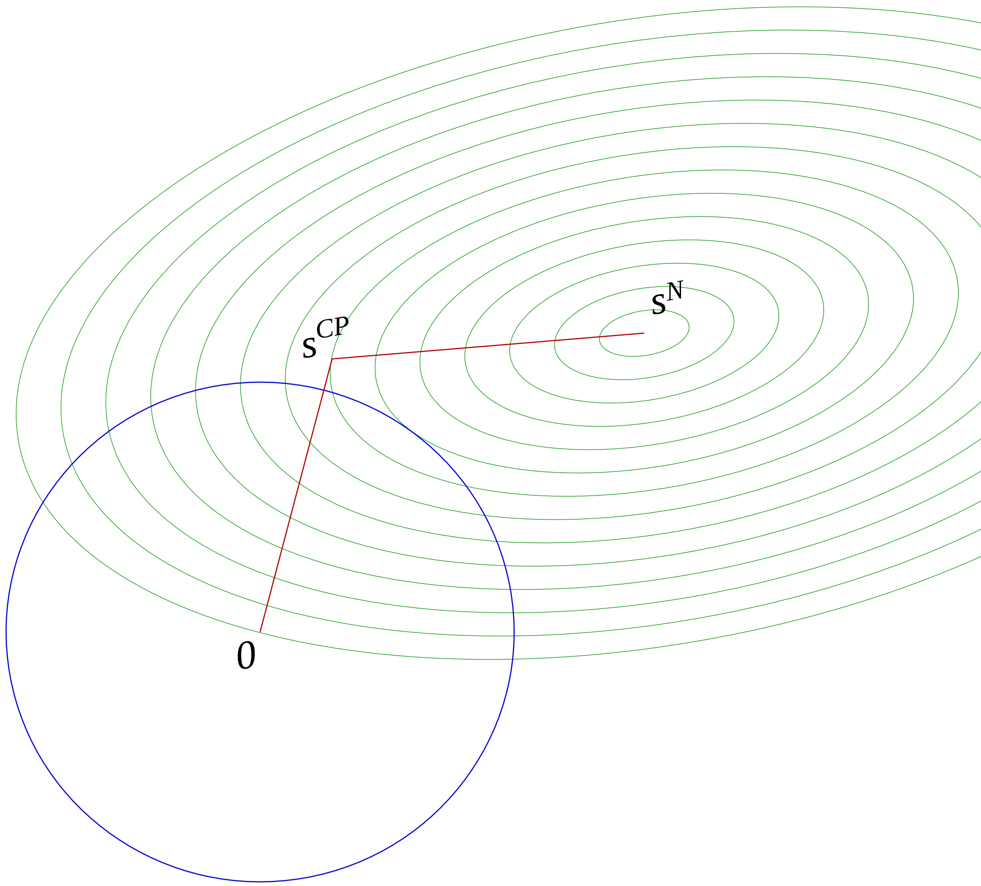
- Can't be computed exactly.



The dogleg step.

- Γ^{DL}: $0 \to s^{\mathrm{CP}} \to s^{N}$.

- $s = \underset{\|w\| \leq \delta, w \in \Gamma^{\mathrm{DL}}}{\arg\min} \|F(u) + F'(u)\, w\|$.



Work toward *inexact Newton* and *Newton–Krylov* adaptations.

Straightforward:

- $\|F(u) + F'(u)\, s^{IN}\| \le \eta \|F(u)\|$
- $\Gamma^{\rm DL}$: $0 \to s^{\rm CP} \to s^{IN}$.

Inexact Newton Dogleg Method:

Given $\eta_{\max} \in [0,1)$, $\delta_{\min} > 0$, $t \in (0,1)$, $0 < \theta_{\min} < \theta_{\max} < 1$,
and initial u and $\delta \geq \delta_{\min}$.

Iterate:

Choose $\eta \in [0, \eta_{\max}]$ and s^{IN} such that

$$\|F(u) + F'(u)\, s^{IN}\| \leq \eta \|F(u)\|.$$

Determine $s \in \Gamma^{\mathrm{DL}}$.

While $ared < t \cdot pred$ do:

Choose $\theta \in [\theta_{\min}, \theta_{\max}]$.

Update $\delta \longleftarrow \max\{\theta\delta, \delta_{\min}\}$.

Redetermine $s \in \Gamma^{\mathrm{DL}}$.

Update $u \leftarrow u + s$ and update δ.

- $ared \equiv \|F(u)\| - \|F(u+s)\|, \quad pred \equiv \|F(u)\| - \|F(u) + F'(u)\, s\|$.
- Choose θ, update δ a la Dennis–Schnabel.
- Determine $s \in \Gamma^{\mathrm{DL}}$ so that $\|s\| \geq \min\{\|s^{IN}\|, \delta_{\min}\}$.

Recall: u is a *stationary point* of $\|F\| \iff \|F(u)\| \leq \|F(u) + F'(u)\, s\| \ \forall s$.

Theorem: *Assume F is continuously differentiable. If u_* is a limit point of $\{u_k\}$, then u_* is a stationary point of $\|F\|$. If additionally $F'(u_*)$ is nonsingular, then $F(u_*) = 0$ and $u_k \to u_*$; furthermore, $s_k = s_k^{IN}$ is acceptable for all sufficiently large k.*

Proof: Since $\|F(u_k) + F'(u_k)\, s_k^{IN}\| \leq \eta_{\max}\|F(u_k)\|$ and $\|s_k\| \geq \min\{\|s_k^{IN}\|, \delta_{\min}\}$, one can show: If u_* is either a non-stationary point or such that $F'(u_*)$ is nonsingular, then there is an $\bar{\eta} < 1$ such that

$$\|F(u_k) + F'(u_k)\, s_k\| \leq \bar{\eta}\|F(u_k)\|$$

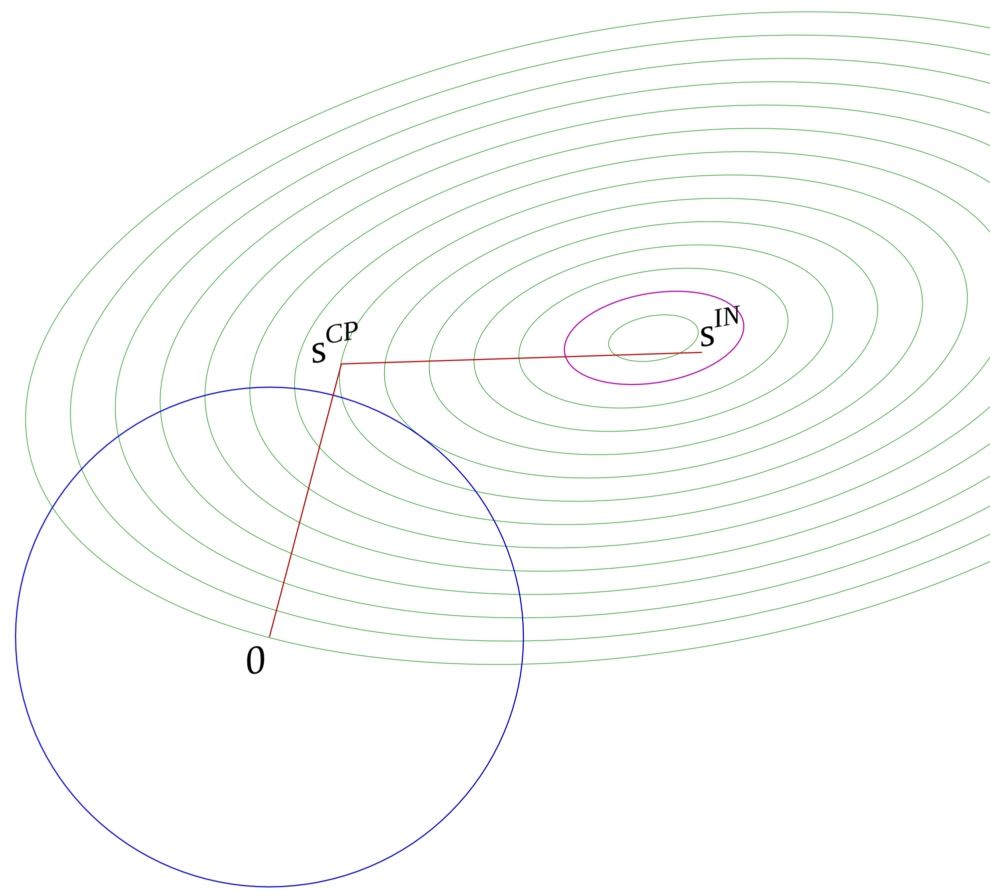
for u_k near u_*. The theorem follows from Eisenstat–W (1994), Cor. 3.6.

Possible **big** problem:

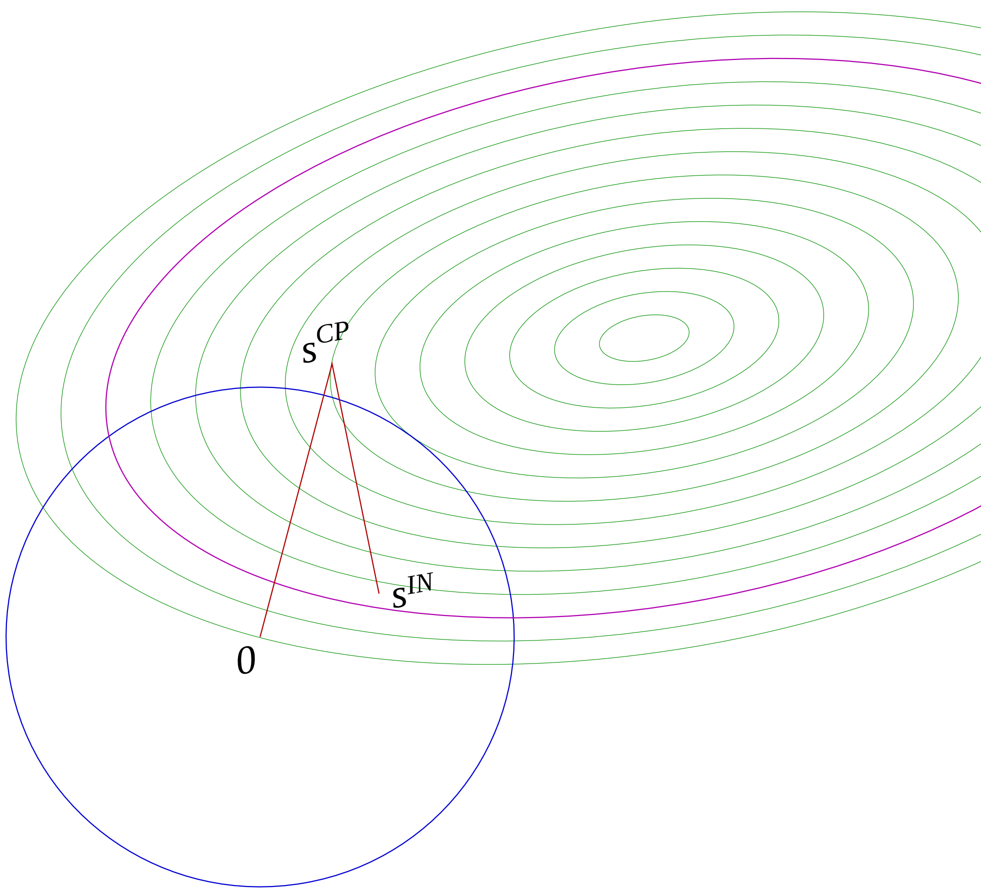
$$\boxed{\textit{Evaluating } s^{\mathrm{CP}} \textit{ requires } F'^T\textit{-products.}}$$

- ▷ Analytic evaluation may be expensive, infeasible.
- ▷ Finite-difference approximation won't work.
- ▷ Automatic differentiation?
- ▷ Brown–Saad (1990): dogleg-within-the-Krylov-subspace using (unrestarted) GMRES.
- ▷ Not a problem when $F' = F'^T$.

Minor consideration: For **any** $\eta \in [0, \eta_{\max})$, $\|F(u) + F'(u)\, s\|$ may not decrease monotonically along Γ^{DL}.



More serious consideration: Unless $\eta \in [0, \eta_{\max})$ is small (how small?), we may have $\langle s^{IN}, s^{\mathrm{CP}} \rangle < \|s^{\mathrm{CP}}\|^2$ or $\|s^{IN}\| < \|s^{\mathrm{CP}}\|$.



How to choose $s \in \Gamma^{\mathrm{DL}}$?

The Standard Strategy.

If $\|s^{IN}\| \le \delta$,

$$s = s^{IN}$$

Else if $\|s^{\mathrm{CP}}\| \ge \delta$,

$$s = (\delta / \|s^{\mathrm{CP}}\|) s^{\mathrm{CP}}$$

Else

$$s = (1-\gamma) s^{\mathrm{CP}} + \gamma s^{IN}$$

for $\gamma \in (0,1)$ such that $\|s\| = \delta$

- $s \in \Gamma^{\mathrm{DL}}$ is uniquely determined.
- s^{IN} is always computed; s^{CP} may not be.
- If η isn't small, we may have $s = s^{IN}$ when $s = \lambda s^{\mathrm{CP}}$ would be preferred.

An Alternative Strategy.

If $\|s^{\mathrm{CP}}\| \geq \delta$,

$$s = (\delta / \|s^{\mathrm{CP}}\|) s^{\mathrm{CP}}$$

Else if $\|F(u) + F'(u)\, s^{\mathrm{CP}}\| \leq \eta \|F(u)\|$,

$$s = s^{\mathrm{CP}}$$

Else if $\|s^{IN}\| \leq \delta$,

$$s = s^{IN}$$

Else

$$s = (1 - \gamma) s^{\mathrm{CP}} + \gamma s^{IN}$$

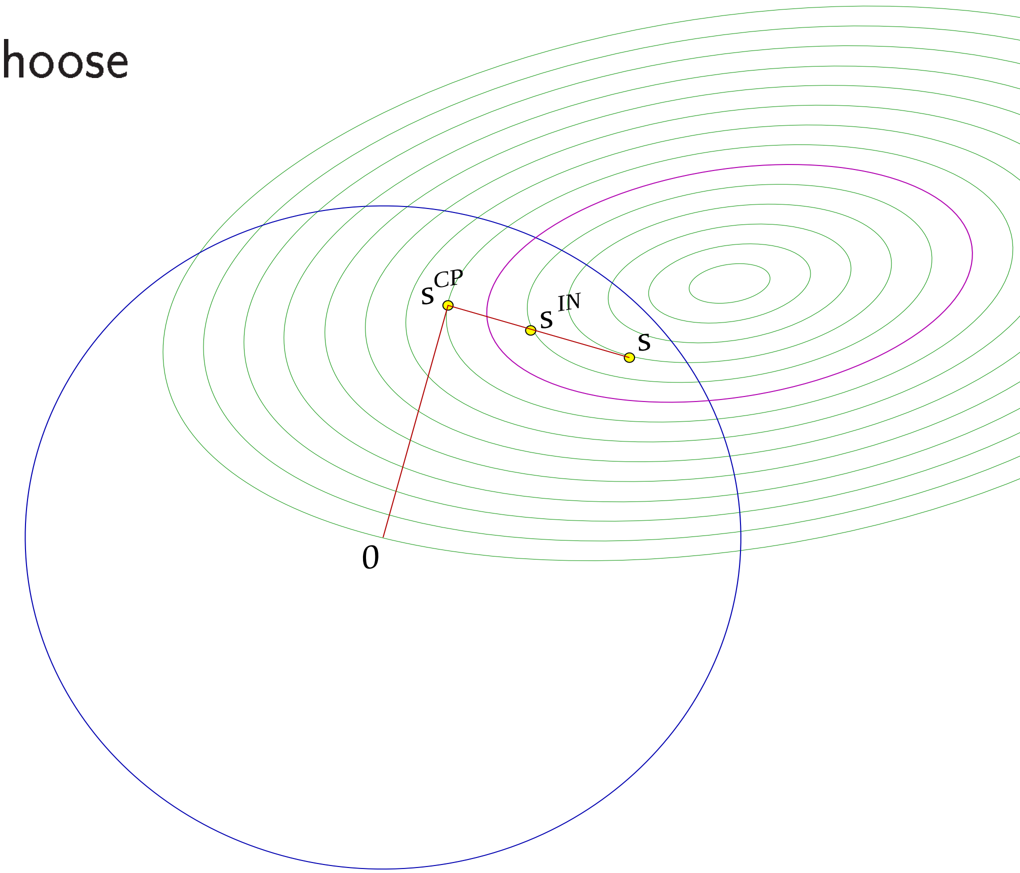
for $\gamma \in (0, 1)$ such that $\|s\| = \delta$

- $s \in \Gamma^{\mathrm{DL}}$ is uniquely determined.
- s^{CP} is always computed; s^{IN} may not be.
- s is appropriately biased toward s^{CP}.

Further refinements.

- If needed, s^{IN} can be computed as $s^{IN} = s^{\mathrm{CP}} + z$, where $\|r^{\mathrm{CP}} + F'(u)\,z\| \le \eta \|F(u)\|$ and $r^{\mathrm{CP}} \equiv F(u) + F'(u)\,s^{\mathrm{CP}}$.

- Having both s^{CP} and s^{IN}, we can choose $s = (1-\gamma)s^{\mathrm{CP}} + \gamma s^{IN}$ so that $\|s\| \le \delta$ and $\|F(u) + F'(u)\,s\|$ is minimal (easy).



Numerical experiments. *Extremely preliminary!!*

- IBM Linux cluster, 4 nodes (8 CPUs).
- MPSalsa + NOX.
- No row-sum scaling (yet).
- Alternative strategy computes $s^{IN} = s^{\mathrm{CP}} + z$, does not minimize $\|F(u) + F'(u)\, s\|$.

2D Thermal Convection Problem. Run times in seconds.

Ra	Backtracking (Quad.)	Dogleg Std.	Dogleg Alt.
10^3	57	56	56
10^4	111	94	93
10^5	146	147	98
10^6	409	1003	265
Geo. Means	139	167	108

Adaptive (Choice 1) Forcing Terms

Ra	Dogleg Std.	Dogleg Alt.
10^3	83	82
10^4	121	126
10^5	293	262
10^6	1266	1171
Geo. Means	247	237

Constant (10^{-4}) Forcing Terms

2D Backward Facing Step Problem. Run times in seconds.

Re	Dogleg Std.	Dogleg Alt.
100	20	23
200	48	36
300	163	30
400	210	35
500	F	63
600	F	137
700	F	F
750	F	F
800	F	F
Geo. Means*	95	28

Adaptive (Choice 1) Forcing Terms

$^{*}100 \leq Re \leq 400$

Re	Dogleg Std.	Dogleg Alt.
100	22	22
200	42	42
300	94	100
400	F	F
500	63	63
600	109	71
700	125	133
750	136	142
800	268	146
Geo. Means**	247	237

Constant (10^{-4}) Forcing Terms

$^{**}Re \neq 400$

Conclusions.

- **None yet!** Except . . .
- These dogleg methods can solve nontrivial problems.
- Methods, strategies, and refinements bear further study.